

# 육각화소 기반의 지역적 이진패턴을 이용한 배경제거 알고리즘

최 영 규<sup>†</sup>

요 약

동영상에서의 배경제거는 다양한 실시간 머신 비전 응용에서 매우 중요한 단계이다. 본 논문에서는 이러한 배경제거를 위한 육각화소 기반의 새로운 접근 방법을 제안한다. 일반적으로 육각형 샘플링 영상은 양자화 오차가 적으며, 이웃화소의 연결성 정의를 크게 개선한다고 알려져 있는데, 제안된 방법은 비매개변수형 배경제거 방법의 하나인 지역적 이진패턴 기반 알고리즘에 이러한 육각 샘플링 영상을 적용하는 것을 특징으로 한다. 이를 통해, 지역적 이진패턴의 추출과정에서 필요한 쌍선형 보간을 없애고 계산량을 줄일 수 있었다. 실험을 통해 이러한 육각화소의 적용이 배경제거 분야에 매우 효율적으로 적용될 수 있음을 확인할 수 있었다.

키워드 : 배경 제거, 움직임 검출, 지역적 이진패턴, 텍스처 분석, 육각화소영상 등

## Background Subtraction Algorithm by Using the Local Binary Pattern Based on Hexagonal Spatial Sampling

Choi, Young Kyu<sup>†</sup>

ABSTRACT

Background subtraction from video data is one of the most important task in various realtime machine vision applications. In this paper, a new scheme for background subtraction based on the hexagonal pixel sampling is proposed. Generally it has been found that hexagonal spatial sampling yields smaller quantization errors and remarkably improves the understanding of connectivity. We try to apply the hexagonally sampled image to the LBP based non-parametric background subtraction algorithm. Our scheme makes it possible to omit the bilinear pixel interpolation step during the local binary pattern generation process, and, consequently, can reduce the computation time. Experimental results revealed that our approach based on hexagonal spatial sampling is very efficient and can be utilized in various background subtraction applications.

Keywords : Background Subtraction, Motion Detection, Local Binary Pattern (LBP), Texture Analysis, Hexagonal Grid Image

### 1. 서 론

영상에서 물체와 배경 영역을 구분하는 것은 지금까지 다양한 머신 비전 응용분야에서 매우 중요한 과정으로 인식되어져 왔다. 특히 최근에는 동영상에서의 배경제거(background subtraction) 요구가 다양한 멀티미디어 응용 분야에서 급격히 증가하고 있다. 예를 들어, 실시간 보안감시(security)에서는 침입자 감지를 위해 배경 영역을 제거해야 하며, MPEG4의 형태기반 코딩에서는 압축률 향상을 위해 영상내의 물체영역의 분리가 필수적이다. 또한 영상검지거나 동작 기반 게임기 등에서도 동영상에서 배경영역을 제거하는 단

계가 전체 시스템의 핵심 기술이 되고 있다. 본 논문에서는 동영상에서 움직이는 객체를 추출하기 위한 배경 제거 방법에 논하고자 한다.

동영상에서 객체의 움직임을 추출하고 배경을 제거하기 위한 목적으로 가장 간단하게 사용되는 방법이 프레임 차이법인데, 이는 인접한 프레임 사이의 차 영상을 만들고, 이를 분석하여 객체영역을 구분한다. 이 방법은 간단하고 계산량이 적다는 장점이 있으나, 잡음의 영향을 매우 많이 받으며, 영상에서 모션을 발생시키는 지역에 대한 정보를 알 수 없는 등의 문제점이 있는데, 이러한 문제를 보완하기 위해서는 후처리 과정이 필요하고 이들에 의한 처리시간이 길어진다는 단점이 있다[1].

이에 비해 배경 차이법은 현재 프레임과 기준이 되는 배경 영상(background image)의 차이를 이용하는 방법으로,

<sup>†</sup> 종신회원 : 한국기술교육대학교 정보기술공학부 부교수  
논문접수 : 2008년 4월 4일  
수정일 : 1차 2008년 7월 28일, 2차 2008년 10월 27일  
심사완료 : 2008년 10월 28일

인접한 두 프레임을 직접 비교하는 것이 아니라 이전 프레임들로부터 배경이라고 할 수 있는 모델(또는 영상)을 유추해 내고 이것과 현재 프레임을 비교하여 배경영역을 제거하는 방법이다. 이 방법은 새로운 프레임을 검사하고 배경모델 갱신을 계속하면서 오래 전 프레임의 영향은 줄이고 현재 프레임의 영향을 추가시키는 방법으로 진행되는데, 시간적 평활법과 시간적 중간치법[2] 등 비교적 간단한 방법들이 많이 사용되고 있다.

Wren등은 이들에 비해 보다 일반화된 방법으로 배경 화소를 하나의 값으로 모델링 하는 것이 아니라 가우시안 분포 (Gaussian distribution)로 가정하고 모델링 하는 방법을 제안하였다[3]. 그러나 이 방법은 정적인 배경에서는 기존의 방법보다 우수하지만 나뭇가지가 움직이거나 잔잔히 물결이 일렁이는 것과 같이 동적인 배경에는 한계가 있었다. 이를 보완하기 위해 각 화소에 대한 모델로 하나가 아니라 여러 개의 가우시안 분포를 허용하는 적응적 가우시안 혼합모델 (Adaptive Gaussian Mixture Model)이 제안되었는데[4], 매 개변수형(parametric) 접근법으로 분류되는 이 방법에서도 기본적으로 화소 값이 가우시안 분포로 나타난다는 가정을 바탕으로 하였다.

그러나 이들은 각 화소마다 여러 개의 분포를 유지하기 위해 많은 처리시간이 소요되어 실시간 응용에 적용하기 어렵고, 또한 화소 값이 가우시안 분포라는 가정이 항상 성립하지는 않는다는 문제가 있는데, 이에 따라 최근에 여러 가지 비매개변수형 (non-parametric) 접근법이 제안되고 있다 [5-8]. 이들은 화소 값의 기반 분포에 대해서 어떠한 가정도 하지 않는데, 이러한 접근법 중에서 최근에 가장 주목을 받는 것이 Pietikainen등이 제안한 지역적 이진패턴(Local Binary Pattern, LBP)을 기반으로 한 배경 모델링 방법이다 [7-8]. 본 논문에서는 최근 좋은 성능을 보고하고 있는 LBP를 기반으로 한 배경제거 방법에 초점을 맞추고, 이 방법의 성능을 개선하기 위한 새로운 접근방법을 제안한다. 특히 기존의 사각형 그리드 영상에 비해 인접화소의 연결성이 탁월한 육각형 그리드 영상을 배경제거에 이용하는 것이 핵심이 되는데, 이를 통해 LBP 생성에 소요되는 처리시간을 대폭 향상시키고, 안정적인 결과를 나타내는 것을 보인다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 기존의 LBP기반 배경제거 방법에 대해 설명하고, 3장에서 제안된

육각형 그리드 기반의 배경제거 방법을 소개한다. 4장에서 실험결과를 보이고 5장에서 결론을 맺는다.

## 2. 지역적 이진패턴(Local Binary Pattern)을 이용한 텍스처

### 2.1 지역적 이진 패턴(Local Binary Pattern)

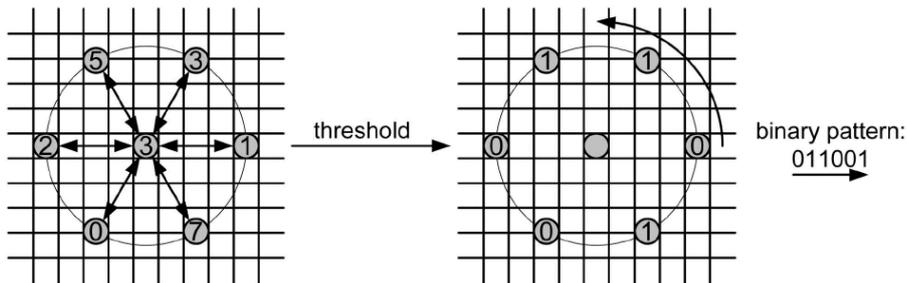
Pietikainen 등은 LBP(Local Binary Pattern)라 불리는 텍스처 분석법을 제안하였다[7]. 이것은 하나의 중앙화소에 대해 이웃화소들과의 관계를 규정하는 일종의 연산자(operator)로, 각각의 이웃화소의 밝기를 중앙화소의 밝기 값으로 이진화하고, 그 결과를 정해진 방향으로 정렬하여 나타내는 것이다. 예를 들어, 이웃화소를 6방향으로 탐색하는 경우, (그림 1)의 왼쪽과 같은 화소 값들이 주어졌을 때 이를 중앙화소 값으로 이진화 하고 이를 반시계 방향으로 탐색하여 코드화하면 이진패턴 011001이 만들어진다. 이 값은 그 중앙화소와 그 이웃에 대한 일종의 명암도 차이 표현으로, 이를 일반화 하면 임의의 화소에 대한 지역적 이진패턴(LBP) 코드는 다음과 같이 정의된다.

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, \quad (1)$$

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

이때,  $g_c$ 는 중앙화소 ( $x_c, y_c$ )의 밝기 값을 나타내며,  $g_p$ 는 중앙화소를 중심으로 반경이 R인 원을 P개로 균등하게 나눈 좌표 ( $x_p, y_p$ )에 대한 밝기 값을 나타낸다. 이때 좌표 ( $x_p, y_p$ )가 정확히 하나의 화소중심에 대응되지 않는 경우 쌍일차 보간(bilinear interpolation)을 통해 밝기 값을 추정하여 사용한다. 이러한 연산자를 통해 P-bit의 패턴이 만들어지게 되며, 따라서 임의의 중앙화소에 대해  $2^P$ 가지의 값들 중 하나가 대응되게 된다.

이러한 LBP 연산의 결과는 중앙 화소에 대한 텍스처 분석에 사용될 수 있는데, 예를 들어, 영상내의 어떤 블록에 대해 그 블록의 모든 화소에서 LBP 코드를 구하고, 이를 히



(그림 1) 지역적 이진패턴(LBP) 계산 방법 [7]

스토그램으로 나타내면 이 히스토그램은 해당 블록에 대한 좋은 텍스처 표현이 된다. 이때 히스토그램의 가로축은 각각의 LBP 코드 값( $0 \sim 2^P - 1$ )이 되며, 세로축은 해당코드가 블록에서 나타난 횟수를 나타내게 되는데, 이러한 히스토그램은 2차원 영상 블록에 대한 에지 특성을 1차원적으로 잘 표현하고 있다고 볼 수 있다. Pietikainen 등은 최근 이러한 방법에 의한 텍스처를 동영상에서의 배경 제거나 얼굴표정 인식 등 다양한 머신비전 응용분야에 적용한 결과들을 활발히 보고하고 있다[7-9].

이와 같은 LBP를 이용한 텍스처 분석 응용에서 텍스처 추출 방향의 수는 계산량과 메모리 사용량에 큰 영향을 미치게 된다. 텍스처 추출 방향의 개수  $P$ 가 너무 크면 가능한 LBP 코드 값의 종류가 많아지고, 이에 따라 히스토그램의 가로축이 지나치게 커지게 된다. 이는 계산량의 급속한 증가를 가져오게 되는데, 예를 들어, 8방향의 LBP를 사용하는 경우  $2^8 = 256$ 가지의 패턴이 나오게 되며 모든 히스토그램의 가로축이 256을 갖는다. 이는 텍스처의 표현력 면에서는 충분하지만 많은 양의 메모리를 필요로 하며 처리시간이 많이 걸리게 되는데, 이것은 이후의 모든 연산이 이러한 히스토그램의 분석과, 비교 및 갱신 등을 기반으로 처리되기 때문이다. 보안감시나 영상검지기 등과 같은 실시간 동영상 응용분야에서는 실시간 처리능력이 어떤 알고리즘의 채택 여부를 결정하는 매우 중요한 요인이 되는데, 텍스처 추출 방향이 너무 많은 경우 실시간 응용에 대한 적용은 어렵게 된다. 또한 실시간 적용이 가능하더라도 하나의 시스템에서 처리할 수 있는 비디오 채널의 수에 따라 시스템의 경쟁력이 좌우되므로 성능에 비해 지나치게 계산량이 많은 방법은 적합하지 않다.

이에 비해, 4 방향으로 LBP를 추출하는 경우, 16가지의 LBP 코드 값이 발생하여 (히스토그램의 명암도 수가 16) 메모리나 계산량 측면에서는 유리하지만, 패턴의 표현력이 떨어지고, 패턴간의 분별력이 지나치게 부족해 진다는 단점이 있다.

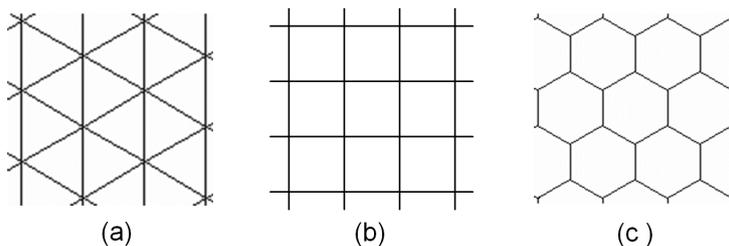
따라서 Pietikainen이 제안한 논문[7]에서는 최적의 값으로 (그림 1)과 같은 6방향의 지역적 이진패턴을 사용하였는데, 이 경우 하나의 화소에 대해 64가지의 LBP 패턴이 발생할 수 있어 메모리나 표현력 측면에서 최적이라고 판단하였다. 본 논문에서 분석해 본 결과에서도 감시영상에서의 배경제거 응용에서는 6방향의 LBP가 계산량과 표현력 면에서 최

적임을 알 수 있었다. 그러나 (그림 1)에서와 같이 6방향의 LBP에서는 모든 이웃화소가 정확히 화소중심에 일치하지 않는다는 문제가 생기는데, 이를 해결하기 위해 쌍일차 보간(bilinear interpolation)이 필요하게 된다. (그림 1)에서와 같이 중심 화소에 대한 6방향의 이웃들 중에서 좌우로 하나씩 두 개의 이웃의 위치는 정확히 화소의 위치에 대응하지만 나머지 4개의 이웃은 정확한 화소위치로 대응되지 않게 되며, 따라서 쌍일차 보간을 통해 근사하게 된다. 사각형 그리드 영상에서의 한 번의 쌍일차 보간에는 3번의 subtraction 연산과 3번의 multiplication 연산이 필요한데, 6방향의 LBP 연산에서는 4방향의 이웃의 위치가 화소 중심과 일치하지 않고, 따라서 각각 12번씩의 연산들이 보간을 위해서 필요하게 된다. 본 논문에서는 이러한 6비트 LBP 코드 생성을 위한 효율적인 방법을 제안하고 이를 배경제거에 적용하는데, 이를 위한 기본 개념은 육각형 그리드의 도입이다.

2.2 육각 샘플링 영상 및 제안된 논문의 동기

Horn은 (그림 2)에서와 같이 평면이 삼각형, 사각형, 및 육각형의 세 가지 도형으로 나뉘어 질 수 있음을 보였는데 [10], 다양한 컴퓨터 비전 응용 분야에서 육각형 그리드가 사각형에 비해 많은 장점이 있음을 설명하였고, 육각 샘플링 영상의 취득이 매우 간단한 소프트웨어나 하드웨어적인 처리를 통해서 이루어질 수 있음을 기술하였다. Sander 등은 영상의 취득시 육각형의 공간 샘플링이 사각형의 샘플링에 비해 양자화 오차(quantization error)가 적다는 것을 증명하였다[11]. 이렇게 영상의 형성 과정에서 육각 샘플링의 사용은 양자화 성능을 향상시키고, 특히 화소의 연결성을 크게 개선하는데, 이에 따라 영상처리나 컴퓨터 그래픽스에서 많이 발생하는 앨리어싱(aliasing) 문제를 개선할 수 있다. (그림 2(c))에서와 같이 육각그리드의 사용은 이웃화소의 정의를 명확히 해 주며, 중앙화소와 이웃화소의 거리를 항상 동일하게 하는데, 이것은 기존의 사각 그리드에서 4-neighbor 와 8-neighbor 선택의 문제를 없앤다. 이러한 특징들을 바탕으로 하여 Cho 등은 사각 그리드에서 적용되었던 에지 이완법(edge relaxation)을 육각 샘플링 그리드에 적용하는 경우 보다 안정적으로 에지를 추출할 수 있다는 것을 보였다[12].

본 논문에서는 동영상에서의 실시간 배경제거 문제에 이러한 육각 샘플링을 적용하고자 하는데, 특히 영상에서  $R=1$  이고  $P=6$ 인 LBP 코드를 추출하는 연산에 적용하고자 한다.



(그림 2) 다양한 평면의 균등 분할 방법: (a) 삼각 그리드, (b) 사각 그리드, (c) 육각 그리드

(그림 2(c))에서 보이는 바와 같이 육각 그리드로 샘플링 된 영상에서 6 방향의 LBP 패턴을 추출하는 경우 모든 이웃의 위치가 정확히 화소중심으로 매핑될 수 있으므로 이웃화소의 추출에서 쌍일차 보간을 적용하지 않아도 좋은 근사를 할 수 있게 된다. 이러한 쌍일차 보간과정의 제거는 LBP 코드 생성을 위한 처리시간을 크게 향상시키고, 결과적으로 LBP를 다양한 실시간 응용에 적용하는데 큰 도움을 준다. 본 논문에서는 Pietikainen[7]이 제안한 LBP를 이용한 배경 모델링에 이러한 육각 샘플링 영상을 적용하고 그 결과를 보인다.

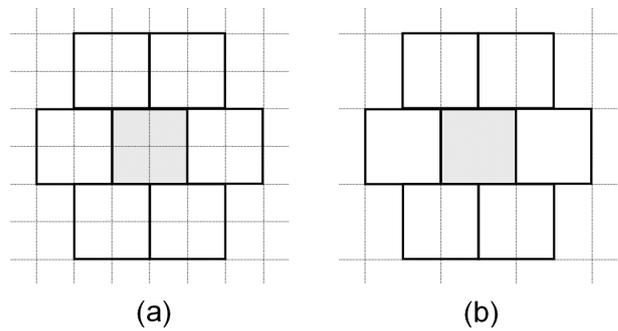
### 3. 육각 그리드 기반의 LBP를 이용한 배경 모델링

배경 모델링(background modeling)은 배경제거(background subtraction)를 위한 가장 중요한 과정으로, 카메라로 취득된 영상에 대한 통계적인 표현을 유지하는 것을 목적으로 한다. 앞에서 소개한 LBP는 이러한 배경모델링에 유익한 여러 가지 특성을 갖는데, 먼저 영상내의 전체적인 명암변화에 영향을 받지 않기 때문에 자연 영상에 대해 밝기의 정규화 과정이 필요 없이 동일한 결과를 나타낼 수 있다. 또한 비매개변수방식(non-parametric) 접근 방법으로 가우시안 분포 등과 같은 배경 분포에 대한 가정이 필요 없고, 처리시간에서 효율적이며, 성능 향상을 위해 최적화해야 할 파라미터가 많지 않은 것으로 알려져 있다. 따라서 본 논문에서는 기반이 되는 배경 모델링 알고리즘으로 Pietikainen 등이 제안한 지역적 이진 패턴 기반 접근법을 채택하였으며, 이 방법에 육각 샘플링 영상을 적용한다. 본 장에서는 입력 영상에서 육각 그리드의 LBP 영상을 생성하는 방법을 설명하고, 이 영상에 배경 모델링 알고리즘을 적용하는 방법을 설명한다.

#### 3.1 육각 그리드에서의 LBP 영상 생성

먼저 육각 샘플링 영상을 구하여야 하는데, 입력 영상이 원래 육각 그리드로 샘플링 된 영상이면 추가적인 과정이 필요 없이 다음 단계로 가서 처리하면 되지만, 그렇지 않은 경우 입력 영상을 육각 샘플링으로 근사되는 영상으로 변환하여야 한다. Horn[10]에 따르면 육각 샘플링은 짝수(even-numbered) 스캔라인에 비하여 홀수(odd-numbered) 스캔라인의 샘플링에 1/2화소만큼의 오프셋을 가함으로써 쉽게 구할 수 있다. 이것은 카메라나 프레임 그레퍼와 같은 기존의 하드웨어를 조금만 수정함으로써 매우 간단히 얻을 수 있다. 특히 TV와 같은 비월주사(interlaced) 방식의 경우 홀수 프레임과 짝수프레임이 나뉘어져 있으므로 더욱 간단히 구현할 수 있다.

그러나 사각형 그리드로 입력된 영상에 대해서는 소프트웨어적으로 먼저 육각 샘플링 영상으로 변환하는 과정이 필요하다. (그림 3)은 사각 그리드 영상을 육각 샘플링 영상의



(그림 3) 육각 그리드 영상의 근사 방법: (a) 영상축소 (b) 반 화소 이동

로 근사하는 방법을 보여주고 있는데, 영상의 축소방법과 반 화소 이동방법을 이용할 수 있다. 먼저 영상 축소는 (a)와 같이 인접한 네 화소 값을 평균하여 축소 영상의 화소를 정의하는데, 축소영상의 홀수 번째 스캔라인의 경우 원 영상에서 한 화소를 오른쪽으로 쉬프트하여 인접한 네 개의 화소를 평균하고, 짝수 번째 스캔라인의 경우 쉬프트하지 않고 평균하여 축소영상의 화소 값을 근사한다. 이에 비해 정밀도를 줄이지 않고 육각 샘플링 영상을 근사하는 방법은 (b)와 같이 홀수 번째 스캔라인의 화소 값을 다시 샘플링하여 구할 수 있는데, 그림의 (b)에서와 같이 모든 인접화소의 중심에서 화소 값을 예측한다. 이때 스캔라인의 두 인접화소를 선형으로 보간하여(linear interpolation) 중심화소 값을 근사할 수도 있으며, 여러개의 인접 화소값을 스플라인 곡선 등으로 모델링하여 화소값을 유추할 수도 있다. 본 논문에서는 320x240 크기의 동영상을 사용하였는데, 배경모델링에서는 정밀도를 반으로 줄여서 160x120크기로 축소하여 알고리즘을 적용하였다.

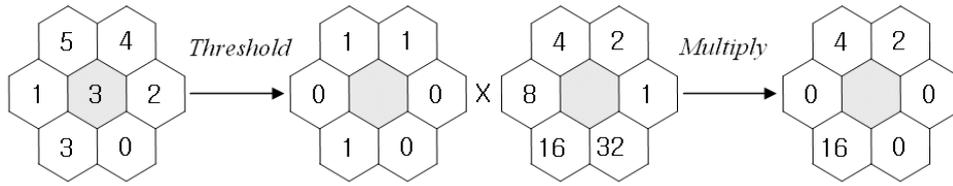
이와 같이 주어진 영상에 대한 육각 샘플링 영상이 만들어지면 배경모델링 단계로 넘어가게 된다.

#### 3.2 LBP를 이용한 배경모델링

Pietikainen 등은 LBP를 제안하고 배경 모델링에 이를 적용하였는데, 앞서 설명한 바와 같이 LBP가 배경 모델링에 여러 가지 유익한 특성을 갖기 때문이다. 배경모델링 과정에서 모든 개별 화소들은 각각 동일한 과정(pixel process)을 통해 처리되는데, 본 논문에서는 하나의 화소에 대한 배경모델링 처리방법에 대해서 기술한다.

먼저 주어진 육각 샘플링 영상에 대한 LBP 영상을 생성하는데, 이를 위해 식 (1)과 (2)가 사용된다. 이때, P=6이며, R=1이다. (그림 4)는 이러한 육각 샘플링 영상의 한 화소(회색의 중앙화소)에 대해 LBP 코드를 생성하는 방법을 잘 보여주고 있는데, 임의의 화소에 대한 LBP의 코드 값은 0~63이 된다.

입력 육각 샘플링 영상에 대한 LBP 영상이 구해지면, 임의의 화소에 대해 먼저 반지름이  $R_{region}$  인 인접 영역에서 LBP 히스토그램을 구하는데, 이런 히스토그램은 중앙 화소



(그림 4) 육각 그리드 영상에서 LBP 코드 생성의 예

와 인접 영역에 대해 1차원으로 표현되는 일종의 텍스처라고 볼 수 있다(히스토그램의 명암도 수(bin수)는 64). 한 화소의 인접 영역에 대한 통계적 특성을 나타내는 이러한 히스토그램을 이용해 그 화소의 배경을 모델링하는 방법이 적용되는데, 프레임이 계속됨에 따라 이들 값들이 갱신되는 적응적인(adaptive) LBP 히스토그램들( $\{m_0, \dots, m_{K-1}\}$ )이 사용된다. 이때 K는 한 화소에서 최대 허용가능한 분포의 수를 나타내며,  $m_k$ 는 k번째 모델 히스토그램을 나타낸다. 하나의 화소에 대한 배경정보로 다수의 LBP 히스토그램을 사용하는 것은 [4]에서와 같이 한 화소에 대한 통계적 표현으로 여러 개의 가우시안 분포를 사용한 것과 동일한 이유인데, 나뭇가지가 바람에 움직이거나 잔잔히 물결이 일렁이는 것과 같이 동적인 배경바람에서 안정적으로 동작하기 위해서이다.

각각의 모델 히스토그램  $m_k$ 는 0에서 1사이의 강도( $\omega_k$ )를 갖는데, 한 화소에 대한 이들 K모델의 강도의 합은 다음과 같이 항상 1로 정규화 된다.

$$\omega_0 + \dots + \omega_{K-1} = 1 \quad (3)$$

새로 입력된 LBP 영상의 해당 화소에서 구한 LBP 히스토그램을  $\vec{h}$ 라 하면, 해당 화소에 대해 현재 저장된 K개의 모델과  $\vec{h}$ 를 비교하여 가장 유사한 히스토그램을 찾게 된다. 이때 히스토그램의 교차(histogram intersection)를 히스토그램의 유사도의 척도로 사용하였다.

$$\bigcap(\vec{a}, \vec{b}) = \sum_{n=0}^{N-1} \min(a_n, b_n), \quad (4)$$

이때,  $\vec{a}$ 와  $\vec{b}$ 는 히스토그램을 의미하고, N은 히스토그램의 전체 명암도의 개수로 6-bit 패턴을 사용하므로 N=64이다. 이 식은 두 히스토그램의 유사한 정도를 표현하고 최대값은 1, 최소값은 0이 되는데, Chi-square이나 Log-likelihood 등의 다른 방법을 적용할 수도 있다.

임의의 화소에서  $\vec{h}$ 와 가장 유사한 모델과의 유사도가  $T_p$  이하이면, 등록된 모델 중 강도가 최저인 모델이  $\vec{h}$ 와 교체되고, 그 화소에 대한 새로운 배경모델로 등록된  $\vec{h}$ 의 초기 강도를  $\omega_{\min}$ 으로 설정한다 (보통  $\omega_{\min}=0.01$ 로 사용됨).

이 경우는 더 이상의 처리가 필요 없으며 그 화소를 물체영역(foreground)로 결정하게 된다.

만약 유사도가 높은 모델이 있는 경우, 새로운 화소는 기존의 배경영역(background)에 속하는 것으로 판단하는데, 이때, 가장 잘 매치가 되는 최적모델  $\vec{m}_k$ 의 히스토그램이  $\vec{h}$ 에 의해 다음과 같이 적응적으로 갱신된다.

$$\vec{m}_k = \alpha_b \vec{h} + (1 - \alpha_b) \vec{m}_k, \quad \alpha_b \in [0, 1] \quad (5)$$

이때,  $\alpha_b$ 는 사용자 지정의 히스토그램의 학습율을 나타내는데, 보통 0.01 내외의 값이 사용된다. 최적모델의 히스토그램 갱신과 함께, 모든 모델에 대한 강도를 다음과 같이 갱신한다.

$$\omega_k = \alpha_\omega M_k + (1 - \alpha_\omega) \omega_k, \quad \alpha_\omega \in [0, 1] \quad (6)$$

이때,  $\alpha_\omega$ 는 분포 강도의 학습율을 나타내며,  $M_k$ 는 최적 모델에 대해서는 1, 그 이외에 대해서는 0이 주어진다. 이것은 가장 잘 매칭된 모델의 경우 강도를 높이고, 그 이외는 강도를 줄여주는 과정이다. 이러한 학습 파라미터들은 모델이 현재의 입력 동영상 시퀀스에 적응하는 속도를 조절해 주는데, 이 값들이 크게 주어지면 입력 영상에 빠르게 적응하게 된다.

마지막으로 임의의 화소에 대해 현재 생성된 모든 히스토그램 모델들 중에서 배경모델로 사용할 분포를 선택하는 과정이 이어진다. 가장 간단한 방법은 개별 모델분포의 강도가 문턱치 이상인 경우 배경으로 사용하는 방법인데, 다음 식과 같이 내림차순으로 정렬된 분포의 합이  $T_B$  이상인 분포들을 배경 히스토그램으로 선택하였다.

$$\omega_0 + \dots + \omega_B \geq T_B, \quad T_B \in [0, 1] \quad (7)$$

이때,  $T_B$ 는 분포 선택을 위한 문턱치로, 움직임이 없는 일정한 (uni-modal) 배경에 대해서는 작은 값을 부여하며, 나뭇잎이 움직이는 등 여러 개의 분포가 필요한 (multi-modal) 배경에 대해서는 보다 큰 값을 지정하게 된다.

물체영역의 검출 (foreground detection)은 배경 모델을 갱신하기 이전에 결정이 된다. 새로운 영상에서 추출한 입

력 히스토그램  $\vec{h}$ 과 이전 프레임까지 추출된 배경 히스토그램들을 식 4를 이용하여 비교해 유사도를 계산하는데, 가장 잘 매칭되는 모델과의 유사도가  $T_p$  이상이면 그 화소는 배경 화소(background)라고 판단되며, 그렇지 않은 경우 물체, 또는 전경 화소(foreground)로 결정된다.

### 3.3 알고리즘의 파라미터 최적화

배경 제거 알고리즘에는 각각의 응용에 대한 성능향상을 위해 다수의 파라미터에 대한 최적화가 필요한데, 이 과정에는 다소 경험적인 방법이 사용된다. 이 절에서는 이들에 파라미터의 설정에 대해 설명한다.

먼저 제안된 방법은 LBP를 추출하기 위한 식 (1)에서의 R과 P는 각각 R=1과 P=6으로 결정이 된다. 이것은 본 논문이 육각 샘플링 영상에 적용을 목표로 하였기 때문이다. 영상의 공간적 중복성 특징에 의해 영상내의 화소간의 상관관계는 거리가 증가함에 따라 감소하고, 가까운 이웃화소들 사이에서 많은 텍스처 정보가 발생한다고 볼 수 있다. 따라서 LBP 추출 시 R값은 보통 적게 유지하는 것이 좋는데[8], 본 논문에서는 R을 1로 고정하였고, 입력의 종류에 따라 영향을 크게 받지 않고 안정적으로 동작하였다. LBP의 방향 P는 R의 영향을 받는데, R이 크면 방향의 수 P도 함께 증가시켜야 한다. P의 값은 또한 메모리의 소비량에 큰 영향을 미치며, 결국 계산량에도 직접적으로 영향을 준다. 히스토그램의 명암도 수(bin수)가  $2^P$ 이므로, 하나의 방향이 증가할 때 마다 메모리와 처리시간이 두 배 가까이 늘어난다고 볼 수 있다. P의 증가가 LBP의 표현량 증가를 가져올 수는 있지만, 이러한 메모리와 처리시간을 증가시키므로, 실시간 처리 응용 분야에서는 적절한 값을 선택하여야 한다. 이들을 고려할 때 육각 샘플링 영상에서 이웃의 관계가 명확한 6방향을 사용하는 것은 표현력과 계산량 측면에서 매우 적절한 선택이 된다.

히스토그램을 계산하는 영역의 넓이를 결정하는  $R_{region}$

은 히스토그램을 보다 지역적으로 추출할 것인지 아닌지를 결정하는데, 원 영상의 크기와 최소 물체영역의 크기 등과 관련이 있으며, 원 영상이 크고 물체영역의 크기가 비교적 큰 경우 보다 큰 값을 사용하고 그렇지 않은 경우 보다 작은 값을 사용한다.

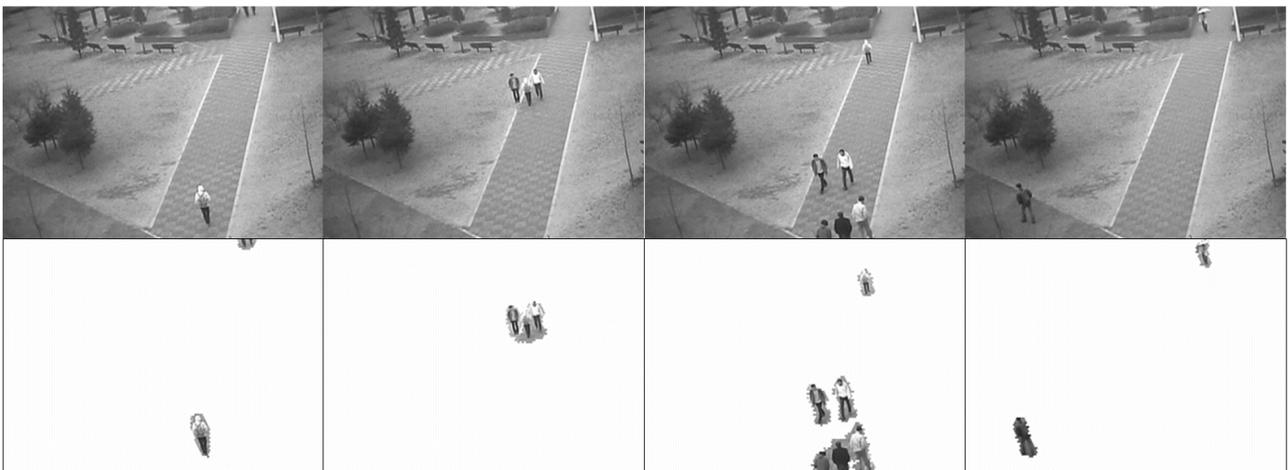
최대 허용 가능한 모델의 히스토그램 수 K는 입력 동영상의 성질에 따라서 영향을 받는데, 바람에 흔들리는 나뭇잎이 많거나, 물결이 일렁이는 등과 같이 비교적 동적인 영상에 대해서는 상대적으로 큰 값을 할당하며, 정적인 동영상에서는 작은 값을 설정하는데, 2~5의 값이 적당하였으며, 이 값이 커지면 계산량도 많아지게 된다.  $T_B$ 도 유사한 성질을 갖게 되는데, unimodal한 영상에서는 작은 값으로도 충분하지만 multimodal한 영상에 대해서는 보다 큰 값으로 설정하게 된다.

히스토그램의 유사도 문턱치  $T_p$ 는 실험적으로 적절히 선택하는데, 실험에 사용된 동영상 시퀀스에 대해 0.6~0.7의 값이 적당한 것으로 나타났다. 모델 갱신의 적응속도 조절하는  $\alpha_b$ 와  $\alpha_w$ 를 큰 값을 적용하면 프레임의 변화에 대한 적응은 빨라지지만 배경 모델의 일관성이 떨어지게 되는데, 실험을 통해 이들은 작은 값으로 유지되는 것이 바람직한 것으로 나타났으며, 0.01 내외의 값이 적절하였다.

## 4. 실험

제안된 방법 Windows-XP 운영체제하의 Pentium-PC에서 C와 C++를 이용하여 구현하였으며 Visual Studio 6.0 환경에서 Release 모드로 컴파일 되었고, 영상의 크기가 320×240인 실내외에서 취득된 여러 비디오 시퀀스에 대해 실험하였다.

(그림 5)는 총 1686 프레임으로 구성되어 있는 “공학1관 앞” 데이터에 대해 제안된 방법을 적용한 결과를 보여주고 있다. 각각의 프레임은 160×120크기의 육각 샘플링 영상으



(그림 5) 공학1관 앞 데이터에 대한 여러 프레임 처리 결과(324, 438, 544, 및 1326프레임)

로 변환하고 LBP 기반의 배경추출 알고리즘을 적용하였는데, 육각 샘플링을 위해 (그림 3)의 (a)와 같은 영상 축소방법을 사용하였다. (그림 6)은 544 프레임의 일부 영역 (a)에 대해 (b)는 사각 샘플링 축소영상을, (c)는 육각 샘플링 축소영상을 보여주고 있는데, 양자화 오차와 이웃화소의 정의에서 육각 샘플링 영상이 사각 샘플링에 비해 우월하다는 [10,11] 것을 그림 (b)와 (c)의 에지영역들에서 짐작해 볼 수 있다. 특히 제안된 방법에서는 이러한 육각 샘플링 적용을 통해 LBP 패턴을 생성할 때 사각 샘플링에서 발생하는 짝일차 보간 과정을 생략할 수 있어 계산량을 대폭 개선할 수 있었다. (d)는 육각 샘플링 영상 (c)에서 추출한 6방향의 LBP 패턴을 보여주는데, 각 화소의 값들이 0~63의 패턴 값을 갖게 되고, 따라서 모든 모델의 히스토그램 명암도 수

(bin수)는 64가 된다. 각 화소에 대해 이러한 히스토그램을 이용해 3.2절에서 소개한 방법을 통해 배경을 모델링하고, 물체영역을 추출해 내는데, (그림 5)의 아래쪽 영상들이 배경이 제거된 결과를 나타내고 있다. <표 1>은 이때 사용되었던 각 파라미터의 값들을 요약해서 보여주고 있는데, 본 논문에서 실험한 모든 데이터에 대해서 동일한 파라미터로 좋은 결과를 나타내는 것을 알 수 있었다.

(그림 7)과 <표 2>는 제안된 방법(P=6)과 P=4 및 P=8로 설정한 경우에 대해 임의의 프레임에 대한 배경제거 실험결과를 비교하여 보여주고 있다. 프레임에 따라 차이는 있지만, 전반적으로 8방향과 6 방향을 사용한 경우가 4 방향을 사용한 경우에 비해 훨씬 배경의 분리가 안정적으로 이루어지는 것을 알 수 있었는데, 이는 4방향의 LBP를 사용하는 경우 가능



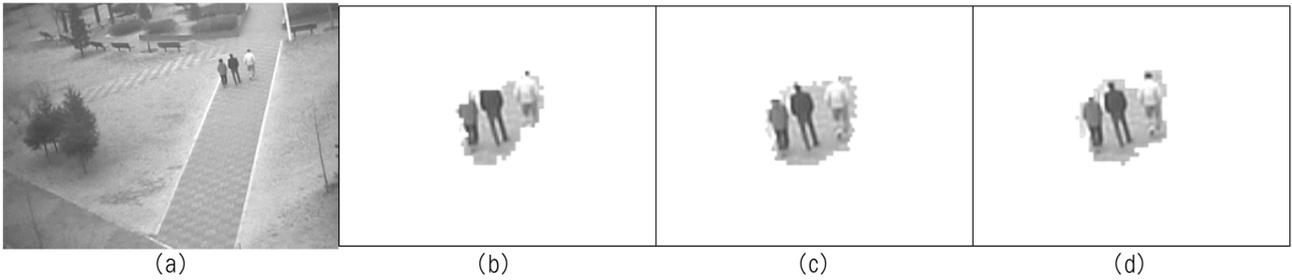
(그림 6) 공학1관 앞 데이터의 544프레임 일부 영역: (a) 원 영상, (b) 사각 샘플링 축소영상, (c)육각 샘플링 축소영상, (d) 육각 축소영상의 LBP 패턴

<표 1> 실험에 사용된 파라미터

Test Sequence	파라미터							
	R	P	$R_{region}$	K	$\alpha_b$	$\alpha_w$	$T_D$	$T_B$
모든 테스트 데이터	1	6	7	3	0.01	0.01	0.65	0.08

<표 2> 프레임 당 평균 처리시간과 주요 메모리 사용량 요약

	프레임별 평균 처리시간(msec)			메모리	
	LBP영상 생성	배경제거 알고리즘	전체 처리시간	NBins (수)	총 메모리 사용량(KB)
$LBP_{4,1}$ (Sampling)	0.36	8.84	9.20	16	2,803
$LBP_{6,1}$ (Bilinear Interpolation)	6.20	16.38	22.58	64	10,176
$LBP_{6,1}$ (Hexagonal Sampling)	0.5	16.36	16.87	64	10,176
$LBP_{8,1}$ (Bilinear Interpolation)	6.44	41.52	47.98	256	39,667



(그림 7) (a) 787번째 프레임의 원영상, (b) 4방향 LBP, (c) 6방향 LBP 및 (d) 8방향 LBP를 이용한 물체영역 추출 결과

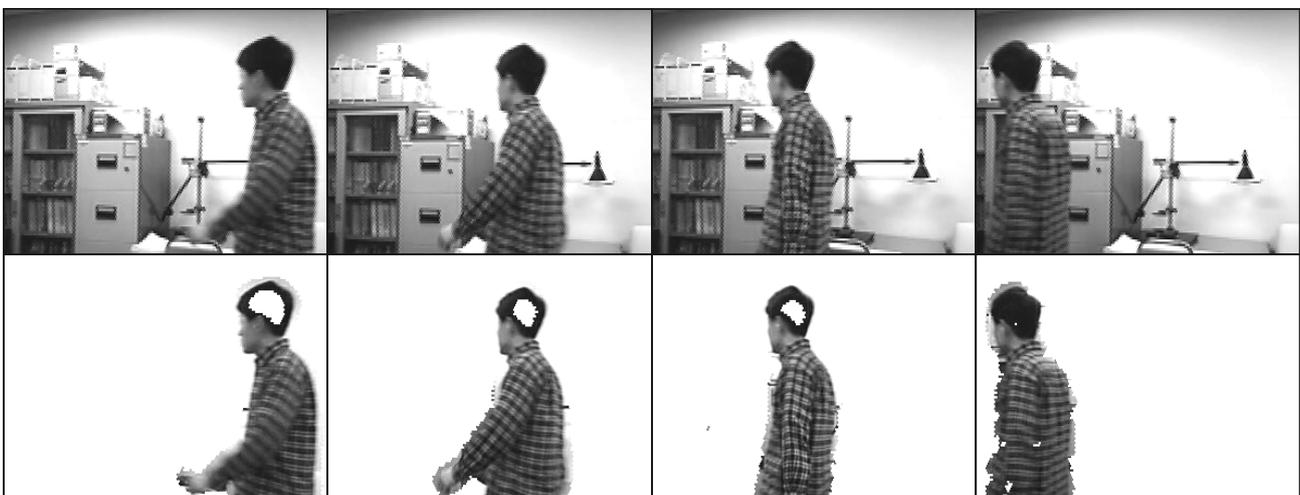
한 패턴의 수가 16이 되어 표현력이 크게 떨어지기 때문으로, 물체영역의 경계영역에서 이런 부분들이 많이 발생하였다. 8방향의 LBP 패턴을 사용하는 경우는 배경추출 결과는 비교적 좋게 나타나지만, 히스토그램의 명암도 수의 증가(256)에 따라 메모리와 처리시간이 급속히 증가하여 실시간 처리에 큰 부담이 되는 것을 알 수 있었다. 결국 Pietikainen

등이 사용한 것과 같이 6방향의 LBP패턴이 실시간 배경제거 응용에 최적임을 확인할 수 있었는데, 대부분의 프레임에서 P=8인 경우와 배경분리 결과는 큰 차이가 없지만, 처리시간은 대폭 개선되는 것을 알 수 있었다. <표 2>는 각각에 대한 처리시간과 메모리 사용량을 요약해서 보여주고 있다.

(그림 8)은 아파트 주차장 데이터에 대한 처리 결과를 보여



(그림 8) 아파트 주차장 데이터에 대한 여러 프레임 처리 결과 (119, 251, 345, 및 403프레임)



(그림 9) 연구실 안 데이터에 대한 여러 프레임 처리 결과 (119, 251, 345, 및 403프레임)

주고 있는데, 두 번째 영상에서 볼 수 있듯이 객체의 크기에도 안정적이며 주차장 바닥의 페인트 색과 비슷한 흰색 차량의 움직임도 강건하게 동작하는 것을 알 수 있다. 마지막으로 (그림 9)는 실험실 내에서 추출한 데이터에 대해 처리한 결과를 보여주고 있는데, 배경을 잘 모델링 하고 안정적으로 객체의 영역을 분리해 낸다. 이 데이터에서 머리의 가운데 부분을 배경영역이라 인식하는 경향이 있었는데, 이것은 이 부분의 밝기 값이 거의 0에 가까워 LBP를 통해서도 유효한 텍스처가 추출되지 않았기 때문이며, 밝기차이가 조금만 더 있더라도(예를 들어 네 번째 영상) 이런 부분이 많이 제거되는 것으로 나타났다. 이러한 실험을 통해 제안된 방법이 지역적 이진패턴 코드를 매우 효율적으로 추출할 수 있으며, 이를 동영상에서의 배경제거 응용에 적용한 경우 매우 안정적이고 효과적인 결과를 나타내는 것을 알 수 있었다.

### 5. 결 론

본 논문에서는 동영상을 이용한 다양한 실시간 머신 비전 응용에서 매우 중요한 단계의 하나인 배경제거(background subtraction)를 위한 새로운 접근방법을 제안하였다. 영상검지기나 동작기반 게임 등과 같은 실시간 응용에서 초당 30프레임의 실시간성의 보장과 시스템당 처리 가능한 채널수 향상을 위한 계산량의 최적화는 매우 중요한데, 본 논문에서는 기존의 사각화소가 아니라 육각화소 영상을 이러한 문제에 적용하였다. 육각형 샘플링 영상은 양자화 오차가 적으며, 이웃화소의 정의에 있어서도 기존의 사각 그리드에 비해 우월하다는 특징이 있는데, 기존 하드웨어의 간단한 수정만을 통해서 쉽게 추출할 수 있다. 제안된 방법에서는 입력 영상을 육각 샘플링 영상으로 변환하고, 이를 비매개변수형 접근법의 하나인 지역적 이진패턴 기반 알고리즘에 적용하였다. 이 과정에서 LBP 패턴 생성에 필요한 쌍일차 보간을 제거하여 계산량을 크게 줄일 수 있었다. 실험을 통해 제안된 방법이 효과적이고 안정적으로 동작하는 것을 알 수 있었는데, 이러한 접근방법이 지역적 이진패턴을 이용한 다양한 텍스처 분석응용에서 활용될 수 있을 것으로 기대된다.

### 참 고 문 헌

[1] 조태훈, 최영규, "다중 배경 분포를 이용한 움직임 검출", 정보처리학회논문지, 8권 4호, pp.381-389, 2001.

[2] N. McFarlane and C. Schofield, "Segmentation and Tracking of Piglets in Images," Machine Vision Application, Vol.8, pp.187-193, 1995.

[3] C. Wren, A. Azarbayejani, T. Darrell and A. Pentland, "Pfinder: Real-Time Tracking of the Human Body," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.19, No.7, pp.780-785, July, 1997.

[4] C. Stauffer and W. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, Vol.2, pp.246-252, 1999.

[5] A. Elgammal, R. Duraiswami, D. Harwood and L.S. Davis, "Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance," Proc. IEEE, Vol.90, No.7, pp.1151-1163, 2002.

[6] K. Kim, T. Chalidabhongse, D. Harwood and L. Davis, "Background Modeling and Subtraction by Codebook Construction," Proc. IEEE International Conf. Image Processing, Vol.5, pp.3061-3064, 2004.

[7] M. Heikkila, M. Pietikainen and J. Heikkila, "A Texture-Based Method for Detecting Moving Objects," Proc. British Machine Vision Conf., Vol.1, pp.187-196, 2004.

[8] M. Heikkila and M. Pietikainen, "A Texture-Based Method for Modeling the Background and Detecting Moving Objects," IEEE Trans. on PAMI, Vol.28, No.4, pp.657-662, April, 2006.

[9] G. Zhao and M. Pietikäinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," IEEE Trans. on PAMI, Vol.29, No.6, pp.915-928, 2007.

[10] B. Horn, "Robot Vision," MIT Press, Cambridge, MA, USA, 1986.

[11] B. Camgar-Parsi and W. Sander, "Quantization error in spatial sampling: comparison between square and hexagonal pixels," Proc. CVPR, 1989, pp.604-611.

[12] T. I. Cho and K. H. Park, "Hexagonal edge relaxation," Electronics Letters, Vol.28, No.4, pp.357-358, 1992.

## 최영규



e-mail : ykchoi@kut.ac.kr

1989년 경북대학교 전자공학과(학사)

1991년 KAIST 전기 및 전자공학과(공학 석사)

1995년 KAIST 전기 및 전자공학과(공학 박사)

1995년~1999년 LG산전연구소 영상연구실 선임연구원

1999년~현재 한국기술교육대학교 정보기술공학부 부교수

관심분야 : 영상 처리 및 인식, 그래픽스, 모델링, 가상현실 등